

Adrian, Tobias; Moench, Emanuel

**Working Paper**

## Pricing the term structure with linear regressions

Staff Report, No. 340

**Provided in Cooperation with:**

Federal Reserve Bank of New York

*Suggested Citation:* Adrian, Tobias; Moench, Emanuel (2008) : Pricing the term structure with linear regressions, Staff Report, No. 340, Federal Reserve Bank of New York, New York, NY

This Version is available at:

<https://hdl.handle.net/10419/60737>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

Federal Reserve Bank of New York  
Staff Reports

Pricing the Term Structure with Linear Regressions

Tobias Adrian  
Emanuel Moench

Staff Report no. 340  
August 2008  
*Revised April 2010*

This paper presents preliminary findings and is being distributed to economists and other interested readers solely to stimulate discussion and elicit comments. The views expressed in the paper are those of the authors and are not necessarily reflective of views at the Federal Reserve Bank of New York or the Federal Reserve System. Any errors or omissions are the responsibility of the authors.

## **Pricing the Term Structure with Linear Regressions**

Tobias Adrian and Emanuel Moench

*Federal Reserve Bank of New York Staff Reports*, no. 340

August 2008; revised April 2010

JEL classification: G10, G12

### **Abstract**

We show how to price the time series and cross section of zero coupon bonds via ordinary least squares regressions. Our approach allows computationally fast estimation of term structure models with a large number of pricing factors. Even though we do not impose cross-equation restrictions in the estimation, we show that our return regressions generate a term structure of interest rates with small pricing errors compared to commonly reported specifications, both in and out-of-sample.

Key words: term structure of interest rates, fama-macbeth regressions

---

Adrian: Federal Reserve Bank of New York (e-mail: [tobias.adrian@ny.frb.org](mailto:tobias.adrian@ny.frb.org)). Moench: Federal Reserve Bank of New York (e-mail: [emanuel.moench@ny.frb.org](mailto:emanuel.moench@ny.frb.org)). First draft of paper: June 4, 2008. The authors would like to thank Richard Crump, Greg Duffee, Darrell Duffie, Arturo Estrella, Eric Ghysels, Paul Glasserman, Scott Joslin, Don Kim, Monika Piazzesi, Joshua Rosenberg, Jonathan Wright, Hao Wu, and seminar participants at the Federal Reserve Bank of New York, the SITE workshop at Stanford University, and New York University for helpful comments and discussions. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

# 1 Introduction

Affine models of the term structure of interest rates are a popular tool for the analysis of government bond yields. These models typically start with three assumptions: 1) the pricing kernel is exponentially affine in the shocks driving the economy, 2) prices of risk are affine in the state variables, and 3) innovations to state variables and log yield observation errors are conditionally Gaussian and independent (see Dai and Singleton (2000, 2002, 2003), Goldstein and Dufresne (2001), Duffee (2002), Kim and Wright (2005) for examples). These assumptions give rise to log yields that are affine in the state variables, and whose coefficients on the state variables are subject to no-arbitrage constraints across maturities (see Duffie and Kan (1996), Piazzesi (2003), and Singleton (2006) for overviews). Empirically, the affine term structure literature commonly uses maximum likelihood methods to estimate coefficients and pricing factors, thus exploiting the distributional assumptions as well as the no-arbitrage constraints.

In this paper, we propose an alternative, regression based approach to pricing the time series and cross section of interest rates. We start with observable pricing factors and develop a three step ordinary least squares estimator. In the first step, we decompose pricing factors into predictable components and factor innovations by regressing factors on their lagged levels. In the second step, we estimate exposures of Treasury *returns* with respect to lagged levels of pricing factors and pricing factor innovations (the latter obtained in the first step). In the third step, we decompose the exposures of returns on the lagged pricing factors (from the second step) into “quasi” prices of risk and yield factor loadings via means of a cross-sectional regression. We compute standard errors that adjust for the generated regressor uncertainty from the previous stages via bootstrap.

Our regression based pricing approach does not rely on distributional assumptions. However, in order to relate the “quasi” prices of risk to actual prices of risk, we need to make distributional assumptions. In particular, we assume that prices of risk are affine

functions of the state variables, that the pricing kernel is exponentially affine, and that shocks to the state variables are conditionally Gaussian. In the appendix, we present an alternative approach with a linear pricing kernel that does not require distributional assumptions.

We treat pricing factors as observable. In our preferred specification, we use the first five principal components of the yield curve as pricing factors. We find that a smaller number of yield factors is not able to adequately capture the dynamics of both returns and yields. In our preferred five factor specification, the Cochrane and Piazzesi (2005) return forecasting factor does not add significant forecasting power for excess Treasury returns.

Our starting point of using yield principal components as factors is not essential. Indeed, we can alternatively use forward rates or the parameters of the Nelson-Siegel-Svensson curve from which our zero-coupon term structure is constructed as pricing factors. The choice of pricing factors does not seem essential as long as the factors span both the time-series *and* cross-sectional variation of the term structure of interest rates.

Our estimation procedure is computationally extremely fast, as it only relies on linear regressions. It is thus easy to use the model for out-of-sample forecasts. We show that our preferred factor specification gives rise to smaller out-of sample pricing errors than A) the random walk model, B) a three factor specification, C) the Diebold and Li (2006) model. Our model can thus be readily used in real time analysis.

Our paper is organized as follows. In Section 2, we discuss our model and the three step estimator. In Section 3, we present our main empirical findings which include an in-sample and out-of-sample analysis of specifications of our model with different sets of pricing factors and at different sampling frequencies. Section 4 concludes.

## 2 The Model

### 2.1 State variables and expected returns

We assume that the dynamics of a  $K \times 1$  vector of state variables  $X_t$  evolve according to the following vector autoregressive process:

$$X_{t+1} = \mu + \Phi X_t + v_{t+1} \quad (1)$$

This specification of the dynamic evolution of the state variables can be interpreted as a discrete time analog to the state variable dynamics of Merton's (1973) ICAPM or Cox, Ingersoll, and Ross' (1985) general equilibrium setup. At this point, we do not make any distributional assumptions about the shocks  $v_{t+1}$ . We solely assume that the expectation of  $v_{t+1}$  conditional on the history of  $X_t$  is zero, and that the variance of  $v_{t+1}$  conditional on the history of  $X_t$  is finite:

$$E_t(v_{t+1} | \{X_s\}_{s=0}^t) = E_t(v_{t+1} | X_t) = 0 \quad (2)$$

$$Var_t(v_{t+1} | \{X_s\}_{s=0}^t) = Var_t(v_{t+1} | X_t) = \Sigma_t \quad (3)$$

where  $\{X_s\}_{s=0}^t$  denotes the history of  $X_t$ . We denote  $rx_{t+1}^{(n-1)}$  the log excess one period holding return of a bond maturing in  $n$  periods. We assume that the return generating process for log excess holding period returns are:

$$rx_{t+1}^{(n-1)} = \underbrace{\beta^{(n-1)'}(\gamma_0 + \gamma_1 X_t)}_{\text{Expected Return}} + \underbrace{\beta^{(n-1)'} v_{t+1}}_{\text{Priced Return Innovation}} + \underbrace{e_{t+1}^{(n-1)}}_{\text{Return Error}} \quad (4)$$

The excess return thus depends on the expected return, a component that is correlated with the innovations of the state variables,  $v_{t+1} = X_{t+1} - E_t[X_{t+1}]$ , and a return pricing error  $e_{t+1}^{(n-1)}$  that is orthogonal to the state innovations. Therefore, the innovations to the state variables are cross-sectional pricing factors, and the levels of the states are forecasting variables. Per construction, the return pricing error  $e_{t+1}^{(n-1)}$  is orthogonal to current and lagged state variables.

Note that up to a convexity term this return generating process is identical to the return generating process that would be implied by a Gaussian affine term structure model that imposes no-arbitrage restrictions. A dynamic term structure model with a linear pricing kernel specification would exactly imply a return generating process as in equation (4). We will discuss the link to arbitrage-free pricing in Section 2.6 below.

## 2.2 Estimation

Based on equation (4) above, we propose the following three step estimator for the parameters of our return equation.

1. Estimate  $\mu$  and  $\Phi$  via OLS. This allows the decomposition of  $X_{t+1}$  into a predictable component  $\hat{\mu} + \hat{\Phi}X_t$ , and an innovation  $\hat{v}_{t+1}$ .
2. Regress excess returns on lagged pricing factors and contemporaneous pricing factor innovations,

$$rx_{t+1}^{(n-1)} = a^{(n-1)} + \beta^{(n-1)'} \hat{v}_{t+1} + c^{(n-1)'} X_t + e_{t+1}^{(n-1)} \quad (5)$$

and recover the coefficients  $\hat{a}^{(n-1)}$ ,  $\hat{\beta}^{(n-1)}$ , and  $\hat{c}^{(n-1)}$ . Stacking them across maturities gives  $\hat{\mathbf{a}} = (\hat{a}^{(1)}, \dots, \hat{a}^{(N)})$ ,  $\hat{\beta} = \left( \hat{\beta}^{(1)'}, \dots, \hat{\beta}^{(N)'} \right)'$ , and  $\hat{\mathbf{c}} = (\hat{c}^{(1)'}, \dots, \hat{c}^{(N)'})'$ , where  $\hat{\mathbf{a}}$  is a vector of length  $N$ , and where  $\hat{\mathbf{b}}$  and  $\hat{\mathbf{c}}$  are matrices of dimension  $N \times K$ , respectively.

3. Estimate the quasi prices of risk parameters  $\gamma_0$  and  $\gamma_1$  via a cross sectional regression.

We know from (4) that  $a^{(n)} = \beta_0^{(n)'} \gamma_0$  and  $c^{(n)'} = \beta_1^{(n)'} \gamma_1 \forall n$ . We therefore have

$\hat{\mathbf{a}} = \hat{\beta} \gamma_0$  and  $\hat{\mathbf{c}} = \hat{\beta} \gamma_1$ . Hence, we can obtain estimates for  $\gamma_0$  and  $\gamma_1$  via

$$\hat{\gamma}_0 = \left( \hat{\beta}' \hat{\beta} \right)^{-1} \hat{\beta}' \hat{\mathbf{a}} \quad (6)$$

$$\hat{\gamma}_1 = \left( \hat{\beta}' \hat{\beta} \right)^{-1} \hat{\beta}' \hat{\mathbf{c}} \quad (7)$$

Note that when  $N = K$ , these estimators amount to setting  $\hat{\gamma}_0 = \hat{\beta}^{-1} \hat{\mathbf{a}}$  and  $\hat{\gamma}_1 = \hat{\beta}^{-1} \hat{\mathbf{c}}$ .

## 2.3 Relation to Fama-MacBeth regressions

The last step of our three step OLS estimation amounts to cross-sectional regressions of  $\mathbf{a}$  on  $\beta$  and  $\mathbf{c}$  on  $\beta$ . Our approach thus bears a close relationship to the cross-sectional regressions of Fama and MacBeth (1973). In contrast to Fama-MacBeth regressions, we assume a particular affine form for the quasi prices of risk. Alternatively, we could follow Fama-MacBeth (1973) by regressing returns  $R_{t+1}$  on  $\beta'$  in the cross-section for each  $t$  to recover  $\gamma_{t+1}^{FM}$ . In our setting, the return generating process implies that  $\gamma_{t+1}^{FM} = (\gamma_0 + \gamma_1 X_t + v_{t+1})$ . We could thus recover  $\gamma_0$  and  $\gamma_1$  from the Fama-MacBeth prices of risk by regressing them on the lagged state variables and a constant. Our approach can thus be interpreted as a dynamic version of Fama-MacBeth regressions that decomposes prices of risk into a predictable component  $(\gamma_0 + \gamma_1 X_t)$  and an innovation  $v_{t+1}$ . The dynamics of prices of risk are in turn linked to the dynamics of the state variables.

## 2.4 Standard errors

Our three step estimation approach uses generated regressors whose estimation uncertainty has to be taken into account in computing standard errors. We use bootstrapping



methods to adjust standard errors for the two layers of generated regressor uncertainty by using the following procedure. We save the residuals  $\{\hat{v}_t\}$  from the VAR(1) regression of the pricing factors on their lagged levels, and the residuals  $\{\hat{e}_t\}$  from the regression of excess returns on the lagged pricing factors and their innovations. Using the stationary block bootstrap of Politis and Romano (1994a,b) which is robust to serial correlation, we then generate 1000 artificial samples of the state vector and the cross-section of returns. In each bootstrap iteration, we simulate the state equation (1) and generate a set of artificial excess returns via (4). For each of the 1000 generated samples of excess returns and factors, we then estimate the market price of risk parameters using our three step regression approach and report the average and standard deviations from the generated sample of estimates as parameter estimates and standard errors, respectively. Standard errors for the three step regression approach could alternatively be computed analytically.

## 2.5 Affine yields

We fit our model to returns of non-defaultable zero coupon bonds. In this section, we show how to calculate the term structure of log bond prices. The definition of the log excess one period holding returns is:

$$rx_{t+1}^{(n-1)} = \ln P_{t+1}^{(n-1)} - \ln P_t^{(n)} + \ln P_t^{(1)} \quad (8)$$

where  $P_t^{(1)}$  denotes the price of a zero coupon bond that matures in one period, and  $P_t^{(n)}$  denotes the time  $t$  price of a zero coupon bond with maturity  $(t+n)$ .

We assume that log bond prices are affine in the states  $X_t$  and an error term  $u_t^{(n)}$ :

$$\ln P_t^{(n)} = A_n + B_n' X_t + u_t^{(n)} \quad (9)$$

We will derive the properties of the yield pricing error term  $u_t^{(n)}$  as a function of the return

pricing error  $e_t^{(n)}$ . By replacing (9) into (8), we see:

$$rx_{t+1}^{(n-1)} = A_{n-1} + B'_{n-1}X_{t+1} + u_{t+1}^{(n-1)} - A_n - B'_nX_t - u_t^{(n)} + A_1 + B'_1X_t + u_t^{(1)} \quad (10)$$

Equating this expression for excess returns with the return generating expression (4), we find:

$$\begin{aligned} \beta^{(n-1)'} (\gamma_0 + \gamma_1 X_t + v_{t+1}) + e_{t+1}^{(n-1)} &= A_{n-1} + B'_{n-1} (\mu + \Phi X_t + v_{t+1}) + u_{t+1}^{(n-1)} \\ &\quad - A_n - B'_n X_t - u_t^{(n)} + A_1 + B'_1 X_t + u_t^{(1)} \end{aligned} \quad (11)$$

This equation has to hold state by state. Let  $A_1 = -\delta_0$  and  $B_1 = -\delta_1$ . Matching terms, we obtain a system of recursive linear restrictions for the bond pricing parameters:

$$A_n = A_{n-1} + B'_{n-1} (\mu - \gamma_0) - \delta_0 \quad (12a)$$

$$B'_n = B'_{n-1} (\Phi - \gamma_1) - \delta'_1 \quad (12b)$$

$$A_0 = 0, B'_0 = 0 \quad (12c)$$

$$\beta^{(n)'} = B'_n \quad (12d)$$

We also obtain the following expression for the log bond pricing errors:

$$\underbrace{u_{t+1}^{(n-1)} - u_t^{(n)} + u_t^{(1)}}_{\text{Log Bond Pricing Error}} = \underbrace{e_{t+1}^{(n-1)}}_{\text{Return Pricing Error}} \quad (13)$$

Two remarks are in order. First, note that the derivation of log bond prices is exact, provided that  $\beta^{(n)'} = B'_n$ . The latter restriction is testable. The recursions (12a) and (12b) are linear ODEs, as we assumed that the return generating model is affine in log

excess returns (see equation 4), and that log bond prices are affine in the state variables (see equation 9). Second, notice that expression (13) above implies that log bond pricing errors are cross-sectionally and serially correlated. Because return innovations should have no serial correlation, the serial correlation of the bond pricing errors arises naturally.

## 2.6 Arbitrage Free Pricing

So far, we have not linked the log-return generating process (4) to arbitrage free pricing restrictions. We will make three standard assumptions in order to introduce no arbitrage constraints in our set up. First, innovations to the state variables  $v_{t+1}$  are conditionally Gaussian. Second, the pricing kernel is exponentially affine. Third, the no-arbitrage equations hold without the pricing error  $e_{t+1}^{(n-1)}$ . These assumptions allow us to link our model to the standard results in the recent literature on affine term structure models (see Dai and Singleton (2000, 2002, 2003), Collin-Dufresne and Goldstein (2001), Duffee (2002), Kim and Wright (2005), Piazzesi (2003), and Singleton (2006)).

Denote a pricing kernel by  $M_{t+1}$ , and the holding period return by  $R_{t+1}^{(n)}$ . The assumption of no-arbitrage implies (see Dybvig and Ross (1987)):

$$1 = E_t \left[ M_{t+1} R_{t+1}^{(n)} \right] \quad (14)$$

We assume that the pricing kernel  $M_{t+1}$  is exponentially affine as in Duffee (2002):

$$M_{t+1} = \exp \left( -r_t - \frac{1}{2} \lambda_t' \lambda_t - \lambda_t' \Sigma_t^{-1/2} v_{t+1} \right) \quad (15)$$

$$\lambda_t = \Sigma_t^{-1/2} (\lambda_0 + \lambda_1 X_t), \quad (16)$$

where  $r_t$  denotes the risk-free one-period interest rate. Noting that  $R_{t+1}^{(n)} = \exp(r_{t+1}^{(n)}) = \exp(rx_{t+1}^{(n)} + r_t)$ , replacing (15) and (4) into (14) and using the fact that  $v_{t+1}$  is conditionally

Gaussian gives:

$$\begin{aligned}
1 &= E_t \left[ \exp \left( -\frac{1}{2} \lambda_t' \lambda_t - \lambda_t' \Sigma_t^{-1/2} v_{t+1} + \beta^{(n-1)'} (\gamma_0 + \gamma_1 X_t) + \beta^{(n-1)'} v_{t+1} \right) \right] \\
&= \exp \left( \beta^{(n-1)'} (\gamma_0 + \gamma_1 X_t) - Cov_t \left[ \lambda_t' \Sigma_t^{-1/2} v_{t+1}, \beta^{(n-1)'} v_{t+1} \right] + \frac{1}{2} Var_t [\beta^{(n-1)'} v_{t+1}] \right) \\
&= \exp \left( \beta^{(n-1)'} (\gamma_0 + \gamma_1 X_t) - \beta^{(n-1)'} (\lambda_0 + \lambda_1 X_t) + \frac{1}{2} \beta^{(n-1)'} \Sigma_t \beta^{(n-1)} \right) \quad (17)
\end{aligned}$$

which can be written as:

$$\beta^{(n-1)'} (\gamma_0 + \gamma_1 X_t) = \beta^{(n-1)'} (\lambda_0 + \lambda_1 X_t) - \frac{1}{2} \beta^{(n-1)'} \Sigma_t \beta^{(n-1)} \quad (18)$$

Following the common assumption that the variance-covariance matrix of shocks to the state variables  $v_{t+1}$  is affine in the state variables:

$$vec(\Sigma_t) = S_0 + S_1 X_t \quad (19)$$

we find:

$$\beta^{(n-1)'} (\gamma_0 + \gamma_1 X_t) = \beta^{(n-1)'} (\lambda_0 + \lambda_1 X_t) - \frac{1}{2} (\beta^{(n-1)'} \otimes \beta^{(n-1)'}) (S_0 + S_1 X_t) \quad (20)$$

These equations can be stacked across maturities  $n$ :

$$\beta (\gamma_0 + \gamma_1 X_t) = \beta (\lambda_0 + \lambda_1 X_t) - \frac{1}{2} \beta^* (S_0 + S_1 X_t) \quad (21)$$

where  $\beta$  is a  $N \times K$  matrix of with the  $n$ -th row given by  $\beta^{(n)'}$  and where  $\beta^*$  is a  $N \times K^2$  matrix with  $n$ -th row given by  $(\beta^{(n)'} \otimes \beta^{(n)'})$ . Setting prices of risk to zero therefore

implies:

$$\lambda_0 = 0 \implies \gamma_0 = -\frac{1}{2} (\beta' \beta)^{-1} \beta' \beta^* S_0 \quad (22)$$

$$\lambda_1 = 0 \implies \gamma_1 = -\frac{1}{2} (\beta' \beta)^{-1} \beta' \beta^* S_1 \quad (23)$$

Hence, the coefficients  $\gamma_0$  and  $\gamma_1$  of equation (4) differ from the price of risk parameters  $\lambda_0$  and  $\lambda_1$  by a convexity term. Note that the latter is particular to the Gaussian case. In general, the wedge between  $\gamma$  and  $\lambda$  depends on the difference between holding period returns and log returns, which happens to be proportional to a convexity adjustment in the Gaussian case. We discuss the equivalent risk-adjustment in the case where shocks are not Gaussian in Appendix A.

Setting the market prices of risk parameters  $\lambda_0$  and  $\lambda_1$  to zero in the recursions (12a) and (12b), one can deduce the risk-neutral bond pricing parameters which in turn can be used to compute risk-neutral yields. These are of independent economic interest since they reflect market participants' expected path of future short term interest rates.

### 3 Empirical results

In this section, we provide estimation results from our regression approach for models with different numbers of factors. As baseline examples, we choose a  $K = 3$  and a  $K = 5$  factor specification where the pricing factors are computed as the first  $K$  principal components from 120 yields for maturities  $n = 1, \dots, 120$  months.<sup>1</sup>

We start by providing in-sample estimation results for the different model specifications. We then turn to out-of-sample yield forecasts and document that the five factor

---

<sup>1</sup>The choice of using principal components of yields as pricing factors was made for convenience and is not essential to our model. Unreported results show that specifications using e.g. the time series of parameters of the Nelson-Siegel curve (as in Diebold and Li, 2006) or a selection of forward rates as pricing factors provide qualitatively similar results.

specification of our model predicts yields well out-of-sample relative to benchmark models.

### 3.1 Data

We estimate our model based on the zero-coupon yield data constructed by Gurkaynak, Sack, and Wright (2006).<sup>2</sup> The construction of these data is based on the Nelson-Siegel-Svensson curve, the parameters of which are published along with the estimated zero-coupon curve. We use these parameters to back out the cross-section of yields for maturities  $n = 1, \dots, 120$  months from which we compute the one-month holding period returns.

We estimate the models for the sample period 1986:01-2008:12 and use as inputs excess holding period returns for bonds of maturities  $n = 12, 18, 24, \dots, 120$  months. This provides us with a cross-section of  $N = 19$  maturities for which we have a total of  $T = 276$  observations. Recall that the excess holding period return for a bond of maturity  $n$  is defined as  $R_{t+1}^{(n)} - R_t^F = \frac{P_{t+1}^{(n)}}{P_t^{(n+1)}} - \frac{1}{P_t^{(1)}}$  where  $P_t^{(n)}$  denotes the price in period  $t$  of a bond which matures in  $t + n$ .

### 3.2 In sample estimation

Taken as given the set of pricing factors, we estimate the parameters  $(\Phi, \gamma_0, \gamma_1)$  using our three step estimation approach.<sup>3</sup> We further obtain estimates of the short rate parameters  $\delta_0$  and  $\delta_1$  by regressing the log one-month yield on the pricing factors. We then feed all parameters into the equations (12a) and (12b) to obtain the recursive pricing parameters and use the latter to compute the model-implied yields.

---

<sup>2</sup>We thank the authors for making these data available for download on the website <http://www.federalreserve.gov/Pubs/feds/2006/200628/feds200628.xls>

<sup>3</sup>Note that our pricing factors are principal components extracted from demeaned yields. Hence, they have a sample average of zero by construction. We therefore fix the mean of the state equation (1) to zero and estimate  $\Phi$  but not  $\mu$ .

### 3.2.1 Three Factor Specification

Table 1 below reports the time series properties of the yield pricing errors  $\hat{u}$  implied by the three factor specification of our model. As these results show, the three factor model explains yields of intermediate maturities very well, implying maximum average pricing errors below 3 basis points across the entire maturity spectrum. Table 2 shows the same set of moments for the return pricing errors  $\hat{e}$  implied by the model.

The yield pricing errors, even though they are small, show substantial serial correlation, as predicted by our equation (13). In contrast, the return pricing errors are essentially serially uncorrelated. Figure 1 provides a plot of average yields as observed and fitted by the three factor model. As can be seen from this plot, the three factor specification provides a very good fit of the average yield curve. This finding is underscored by Figure 2, which shows the unconditional standard deviations of observed and model-implied yields. While these almost coincide for intermediate and long maturities, there is some notable difference at the short end of the curve. Figures 3 and 4 show the time series of observed and fitted yields and excess holding period returns for a set of different maturities. These plots provide additional evidence that the three factor specification of our model gives a close fit to the term structure of interest rates, especially so for intermediate and longer maturities.

Recall that we use the first three principal components of the yield data as the pricing factors in the three factor specification of our model. According to their loadings on individual bond yields, these are commonly labeled "level", "slope", and "curvature". Our arbitrage-free term structure model allows us to derive the risk premia associated with shocks to these components. Figure 5 provides a plot of the yield loadings  $-\frac{1}{n}B_n$  implied by our model. Shocks to the first principal component affect yield of all maturities by about the same amount, so it can clearly be interpreted as a level factor. Moreover, the yield loadings of the second and third factor show that these can readily be viewed

as representing slope and curvature. Hence, the three principal components retain their interpretation when used as state variables in our no-arbitrage model.

While the loadings of yields of different maturities on the level, slope, and curvature factors have often been documented in the literature, relatively little attention has previously been given to their impact on expected excess returns. In our arbitrage-free model, this is straightforward to analyze. Figure 6 shows the loadings  $B'_n\gamma_1$  of expected returns on the three factors. According to this plot, the yield curve level and slope factors are equally strong drivers of variations in risk premia, albeit with opposite signs. Indeed, while the impact on expected excess returns is increasing for the slope factor across the maturity spectrum, it is decreasing for the level factor. Compared to level and slope, the curvature factor only has a relatively small impact on expected returns. We will see below how these conclusions change when we allow additional factors to enter the model.

### 3.2.2 Five Factor Specification

We have seen in the previous section that a three factor specification of our pricing model, although implying small pricing errors, still leaves some of the variation of interest rates at the short end of the yield curve unexplained. We now show that a five factor specification of our model fits the yield curve close to perfectly in-sample. Since traditional term structure models are estimated imposing non-linear cross-equation restrictions, estimation of these models with more than three factors becomes computationally very demanding. In contrast, our estimation approach is based on simple linear regressions, and it therefore comes at no cost to add pricing factors to the model.

Table 4 reports the time series properties of the yield pricing errors implied by the five factor specification of our model. The average yield pricing errors are very small, ranging between 0.3 and 0.7 basis points. Moreover, the standard deviation of the pricing errors is very small, exceeding 1 basis point only for the 12-month maturity. Consistent



with our decomposition of yield pricing errors in terms of return pricing errors, we again find evidence for serial correlation in yield pricing errors. Table 5 shows that the return pricing errors, in turn, are essentially not autocorrelated.

Figures 7 to 10 visualize the above results. As can be seen from these plots, the five factor specification provides a very good fit to the yield curve across the maturity spectrum. In sum, the ability of our model to fit yields of various maturities is striking given that all estimates are based on simple linear regressions. One might think that this is not surprising since our pricing factors are the principal components of yields. However, it is important to keep in mind that the model has been fitted to returns and that all parameters have been obtained via linear regressions without imposing the cross-equation restrictions for yields as in standard affine term structure models. These have only been used ex-post to back out the model-implied yield curve.

While the first three principal components of yields have a common interpretation as level, slope, and curvature, higher order principal components of yields have traditionally not been given much attention in the term structure literature as they only explain minimal shares of the cross-sectional variation of yields. In Figure 11, we plot the yield loadings implied by our five factor model. According to this plot and consistent with the in-sample pricing errors, the fourth and fifth principal component carry a very small amount of explanatory power at the very short end of the curve, but explain little to no variation in longer maturities. Turning now to the effects of these two factors on expected excess returns shown in Figure 12, this picture changes substantially. Indeed, shocks to the fourth and fifth principal components are the main driving forces behind movements in risk premia. Indeed, in the five factor specification, the fourth principal component has a strong positive hump-shaped effect on expected one-month excess returns. In contrast, the fifth principal component exhibits strongly negative coefficients on expected excess returns. In spirit, these results are in line with the findings in Cochrane and Piazzesi

(2005) and Duffee (2009a) who document that factors with negligible contemporaneous effects on the yield curve may have strong predictive power for future excess returns.

Recall that we have derived the recursive pricing equations (12a) and (12b) by equating our return generating process (4) with the definition of excess returns (8) under the assumption that log bond prices and thus yields are linear in the state variables  $X$ . The matching of terms implies that our recursive bond pricing parameters  $B_n$  which are functions of the parameters  $\mu, \Phi, \delta_0, \delta_1, \gamma_0$ , and  $\gamma_1$  estimated via linear regressions be equal to the coefficients  $\beta^{(n)}$  from the time series regressions of excess returns on the factor innovations  $v$  for all maturities  $n$ . In practice, however, this is not by construction the case. Indeed, it may serve as an additional diagnostic of model fit whether the imputed recursive pricing parameters  $B$  equal the regression coefficients  $\beta$ .

Figure (13) provides plots of the two sets of coefficients. As expected, they are visually indistinguishable for all maturities for the first three factors. While they differ slightly for the fourth and fifth factor, that difference is economically and statistically negligible.

### 3.3 Out-of-sample forecasts

We have seen above that both a three factor and a five factor specification of our term structure model fit the yield curve very well in-sample. While the additional factors in the five factor specification do not noticeably enhance in-sample fit, they seem to be the key driving force behind model-implied expected excess returns. In this subsection, we thus assess whether the added pricing factors carry additional predictive power.

To this end, we perform the following recursive out-of-sample forecast exercise. We use the subperiod 1986:01-2002:12 as our training sample and recursively re-estimate the model for all months from 2003:01 until 2008:03 using data from 1986:01 to the month when the forecast is made. That is, we extract the first  $K$  principal components from the yields up to date  $t$ , use these as pricing factors, estimate the model parameters with our

three step regression approach, and then predict yields according to

$$y_{t+h|t}^{(n)} = \hat{a}_n + \hat{b}_n \hat{X}_{t+h|t}$$

where  $\hat{a}_n$  and  $\hat{b}_n$  have been back out from (12a) and (12b) based on estimates of the model parameters obtained using data up to date  $t$ , and where  $\hat{X}_{t+h|t}$  is the  $h$ -step ahead conditional forecast of the state variables implied by their VAR(1) representation.

Table 7 reports the root mean squared forecast errors for the yield predictions implied by the three and five factor specifications of our model at forecast horizons 1-month, 6-months, and 12-months ahead. We further provide results obtained from the five factor specification augmented by an estimate of the Cochrane-Piazzesi return forecasting factor. All RMSEs are stated relative to those implied by a simple random walk model for yield levels. As a means of comparison, we also compute yield forecasts based on the dynamic extension of the Nelson-Siegel (1987) model that has been suggested by Diebold and Li (2006). Nelson and Siegel’s model is not arbitrage-free, but simply decomposes variations in the cross-section of yields at every point in time into three factors with pre-specified loadings. Diebold and Li (2006) give a dynamic interpretation to these time-varying fitted parameters, estimate their dynamics with a simple VAR, and show that this generates good out-of-sample yield forecasts.

Let’s first compare the out-of-sample predictive ability at the one-month ahead horizon. The Nelson-Siegel model is outperformed by the random walk as indicated by relative RMSEs that are larger than one. The same is true for the three factor specification of our model for all maturities but the 12-month yield. In contrast, the five factor specification outperforms all other models including the random walk for maturities up to five years. The improvement relative to the RW forecast appears to be relatively small, though.

The improvement over the random walk is considerably larger for forecasts six and

twelve months ahead. Indeed, we find that our five factor specification achieves a reduction in RMSEs with respect to the random walk of the order of 20% for maturities up to five years. This is comparable in magnitude to the reduction in yield forecast errors that can be achieved using principal components extracted from a large cross-section of macroeconomic variables as pricing factors as in Moench (2008). Our model also clearly outperforms the Nelson-Siegel model. We interpret these results as evidence that factors which may have little explanatory power in the cross-section, might well be useful for predicting future yields and hence represent important state variables. This is again in line with the results in Cochrane and Piazzesi (2005) who find that a linear combination of forward rates which only explains a small share of the cross-sectional variation of yields is a strong predictor for one-year excess holding period returns. As a rough check of whether the fourth and fifth principal components span the information in the CP-factor, we also compare the results of our five factor specification to one that is augmented by an estimate of the CP-factor from our data sample.<sup>4</sup> These results show that the CP-factor only marginally reduces forecast errors, and in some cases even slightly deteriorates yield predictions at the long end of the curve. We therefore conclude that the CP-factor does not carry important predictive information beyond the first five principal components of yields in our sample.

What drives the superior forecast performance of our five-factor model? We provide the following interpretation. We have seen in Section 3.2.2 that the fourth and fifth factor in our model are the most important drivers of expected excess holding period returns. Hence, even if these factors are only marginally informative about the contemporaneous cross-sectional variation of yields, they contain useful information for future yield changes or returns. A three factor model, while fitting the cross-section of yields very precisely at

---

<sup>4</sup>We construct the CP factor by regressing log one-year excess holding returns onto the one-year yield, as well as the four and seven year forward rates. This gives a similar tent-shaped pattern of regression coefficient as in Cochrane and Piazzesi. We use the linear dependence to construct the return-forecasting factor.

any given point in time, is therefore likely to provide worse forecasts.

It appears useful to discuss this interpretation with respect to some recent work by Joslin, Singleton, and Zhu (2009), and Duffee (2009b) who argue that no-arbitrage restrictions do not improve out-of-sample forecasts of yields unless restrictions are imposed on the time series properties of model-implied risk premia. The core of the argument by Joslin et al. (2009) is that with a fully flexible affine market price of risk specification, the physical and risk-neutral factor dynamics are not tied to one another. Duffee (2009b) argues that restrictions on the physical dynamics of the risk factors - without imposing no-arbitrage restrictions - improve the predictive power of out-of-sample yield forecasts the most. In particular, he suggests to force long-maturity yields implied by the model to follow random walk processes.

As our estimation methodology is entirely based on linear regressions, imposing any kind of linear restriction on the parameters of our model is easily implemented. Here, we do not consider restrictions on either the physical factor dynamics or the market prices of risk. Instead, we add two factors to the standard three factor specification and find that this improves the model's forecasts considerably. Unreported results underscore our above interpretation that the additional factors help the forecasts of yields by improving the forecasts of the first two principal components of yields. Indeed, out-of-sample forecast errors of our five factor specification for the level and slope factors are considerably lower than out-of-sample forecasts of the three-factor specification. In contrast, out-of-sample forecasts of level and slope implied by the five factor specification augmented with the CP factor are about the same as those implied by the five factor specification. Hence, the CP factor does not add information to the factor forecasts beyond the first five principal components. Following the logic of Joslin et al. (2009) and Duffee (2009b), imposing restrictions on the specification of risk premia or the physical factor dynamics is likely to further improve our results.

### 3.4 Term Structure Estimation in Real Time

As we fit our model using simple linear regressions, estimation is extremely fast. This is in sharp contrast to traditional likelihood based estimation of term structure models subject to nonlinear cross-equation restrictions which typically requires a long optimization process.<sup>5</sup>

Its simplicity and the speed at which it is estimated makes our model particularly appealing for real time analysis of term structure dynamics. In this section, we document how we fit our model to the US Treasury yield curve at the daily frequency. This allows us to interpret yield curve movements in terms of risk-neutral yield versus term premia dynamics in real time.

Estimation at the daily frequency requires a slight modification of the empirical approach outlined above which was tailored to the monthly frequency. In the daily application of our model, we use the daily yields of maturities from  $n = 1, \dots, 120$  months obtained from Gurkaynak, Sack, and Wright (2006). We aggregate the daily yields to the monthly frequency by taking monthly averages and extract principal components from the monthly yields. We then apply the weights from the monthly principal components to the daily yields in order to obtain daily estimates of our pricing factors.<sup>6</sup>

We compute log excess holding period returns at the monthly frequency and estimate the parameters of our model as outlined in Sections 2.2 and 2.5. Finally, we use the estimated parameters and the daily yield factors to impute model-implied yields and term premia at the daily frequency.

---

<sup>5</sup>Joslin, Singleton, and Zhu (2009) provide a new parameterization of standard Gaussian dynamic term structure models which allows them to estimate a subset of the model parameters using OLS and the remaining parameters using MLE. Their estimation is therefore computationally considerably less burdensome than that of standard affine models.

<sup>6</sup>Note that since we extract the principal components from demeaned monthly yields, we need to make an adjustment to the daily factors. Precisely, we apply the monthly principal components weights to the sample average of the monthly yields and then subtract this vector from the daily factors obtained as described before.

To illustrate the ability of our model to fit daily term structures, Figure 14 shows for the  $K = 5$  factor specification the daily observed and model-implied yields for the 5-year Treasury and the 10-year Treasury notes since August 2008. The model prices both maturities very precisely.

Perhaps more interesting is an analysis of term premia over this sample period. As discussed in Section 2.6 above, we can use our model to achieve a decomposition of interest rates into risk-neutral yields and term premia. Figure 15 displays the estimated term premium for the 10-Year Treasury along with the 1-month MOVE index. The latter is a measure of implied volatilities from options on Treasury futures constructed by Merrill Lynch. The plot shows that our term premium estimate and the MOVE index exhibit a strong correlation. This is striking since our model is estimated without using any option data. Note also that the risk premium adjustment is made for the case of a homoskedastic Gaussian model with  $S_0 = \text{vec}(E[\nu'_t \nu_t])$  and  $S_1 = 0$  and thus  $\lambda_1 = 0$  implies  $\gamma_1 = 0$ . We interpret the correlation between the two time series as evidence that the term premium estimate from our model indeed reflects the risk of holding Treasury securities.

## 4 Conclusion

We outline an empirical approach to the estimation of dynamic term structure models. Our approach is computationally fast, gives rise to small pricing errors, and provides very good out-of-sample forecasts compared to benchmark models. Our contribution is readily implementable for other pricing applications—such as in the macro-finance literature—and to other asset classes. Most importantly, our method is practical for real time analysis.

## References

- Campbell, John Y., Lo, Andrew, and A. Craig MacKinley, 1997, *The Econometrics of Financial Markets*, Princeton University Press.
- Cochrane, John, 2005, *Asset Pricing*, Princeton University Press, Revised Edition.
- Cochrane, John, and Monika Piazzesi, 2005, "Bond risk premia," *American Economic Review* 95, pp. 138-160.
- Collin-Dufresne, Pierre, and Robert S. Goldstein, 2001, "Do Bonds Span Fixed Income Markets? Theory and Evidence for 'Unspanned' Stochastic Volatility," *Journal of Finance*, pp. 1685-1730.
- Cox, John C., Ingersoll, Jonathan E., and Stephen A. Ross, 1981, "A Theory of the Term Structure of Interest Rates", *Econometrica* 53(2), pp. 385-407.
- Dai, Qiang, and Kenneth Singleton, 2000, "Specification Analysis of Affine Term Structure Models," *Journal of Finance* 55, pp. 1943-1978.
- Dai, Qiang, and Kenneth Singleton, 2002, "Expectations Puzzle, Time-Varying Risk Premia, and Affine Models of the Term Structure," *Journal of Financial Economics* 63, pp. 415-441.
- Dai, Qiang, and Kenneth Singleton, 2003, "Term Structure Dynamics in Theory and Reality", *Review of Financial Studies* 16, pp. 631-678.
- Diebold, Frank X., and C. Li, 2006, "Forecasting the Term Structure of Government Bond Yields, *Journal of Econometrics* 130, pp. 337-364.
- Duffie, Darrell, and Rui Kan, 1996, "A Yield-Factor Model of Interest Rates," *Mathematical Finance* 6, pp. 379-406.
- Duffee, Gregory R., 2002, "Term Premia and Interest Rate Forecasts in Affine Models," *Journal of Finance* 57, pp. 405-443.
- Duffee, Gregory R., 2009a, "Information in (and not in) the term structure," *mimeo*,



Johns-Hopkins-University.

Duffee, Gregory R., 2009b, "Forecasting with the Term Structure: The Role of No-Arbitrage Restrictions," *mimeo*, Johns-Hopkins-University.

Dybvig, Philip H., and Stephen A. Ross, 1987, "Arbitrage," *The New Palgrave: a Dictionary of Economics*, pp. 100-106.

Evans, Martin, 2003, "Real Risk, Inflation Risk, and the Term Structure," *Economic Journal* 113, pp. 345-389.

Fama, Eugene F., and James D. MacBeth, 1973, "Risk, return, and equilibrium: Empirical tests," *Journal of Political Economy* 81, pp. 607-636.

Gurkaynak, Refet S., Brian Sack, and Jonathan H. Wright, 2006, "The U.S. Treasury Yield Curve: 1961 to the Present," *Journal of Monetary Economics* 54, pp. 2291-2304.

Hamilton, James D., 1994, *Time Series Analysis*, Princeton University Press.

Hansen, Lars-Peter, and Ravi Jagannathan, 1993, "Assessing Specification Errors in Stochastic Discount Factor Models," *Journal of Finance* 48, pp. 1091-1092.

Scott Joslin, Kenneth J. Singleton, and Haoxiang Zhu, 2009, "A New Perspective on Gaussian DTSMs", *mimeo*, Stanford University.

Kim, Don, and Jonathan H. Wright, 2005, "An Arbitrage-Free Three-Factor Term Structure Model and the Recent Behavior of Long-Term Yields and Distant-Horizon Forward Rates," *Board of Governors of the Federal Reserve System Finance and Economics Discussion Series* 2005-33.

Merton, Robert. C., 1973, "An intertemporal asset pricing model," *Econometrica* 41, pp. 867-887.

Moench, Emanuel, 2008, "Forecasting the Yield Curve in a Data-Rich Environment: A No-Arbitrage Factor-Augmented VAR Approach", *Journal of Econometrics* 146(1), 26-43.

Nelson, Charles R. and Andrew F. Siegel, 1987, "Parsimonious Modeling of Yield

Curves”, *Journal of Business* 60(4), 473-489.

Piazzesi, Monika, 2003, ”Affine Term Structure Models,” *Handbook of Financial Econometrics*.

Politis, Dimitris N., Romano, Joseph P., 1994a, ”Limit theorems for weakly dependent hilbert space valued random variables with applications to the stationary bootstrap,” *Statistica Sinica* 4 (2), 461–476.

Politis, Dimitris N., Romano, Joseph P., 1994b, ”The stationary bootstrap,” *Journal of the American Statistical Association* 89 (428), 1303–1313.

Singleton, Kenneth J., 2006, *Empirical Dynamic Asset Pricing*, Princeton University Press.

# Appendix

## A Arbitrage free pricing without Gaussian shocks

In section 2.6, we made the assumption that the pricing kernel is exponentially affine, and that shocks to the state dynamics are conditionally Gaussian. We can alternatively derive the relationship between the quasi prices of risk  $\gamma_0$  and  $\gamma_1$  and actual prices of risk by assuming that the pricing kernel is affine, without making distributional assumptions about the state innovations. Instead of the pricing kernel of equation (15), we make the following assumption about pricing kernel innovations:

$$\frac{M_{t+1} - E_t[M_{t+1}]}{E_t[M_{t+1}]} = -(\lambda_0 + \lambda_1 X_t)' \Sigma_t^{-1} v_{t+1}. \quad (24)$$

We denote excess holding period returns by  $Rx_{t+1}^{(n-1)}$ . Then the no arbitrage constraint is:

$$0 = E_t \left[ M_{t+1} Rx_{t+1}^{(n-1)} \right]. \quad (25)$$

This can be rewritten as:

$$E_t \left[ Rx_{t+1}^{(n-1)} \right] = -Cov_t \left[ Rx_{t+1}^{(n-1)}, \frac{M_{t+1}}{E_t[M_{t+1}]} \right]. \quad (26)$$

We define the difference between excess holding period returns and log excess returns as  $\omega_{t+1}^{(n-1)}$ :

$$\omega_{t+1}^{(n-1)} = Rx_{t+1}^{(n-1)} - rx_{t+1}^{(n-1)}. \quad (27)$$

Replacing this in the no arbitrage equation gives:

$$\begin{aligned}
E_t \left[ r x_{t+1}^{(n-1)} \right] &= E_t \left[ R x_{t+1}^{(n-1)} - \omega_{t+1}^{(n-1)} \right] \\
&= -Cov_t \left[ r x_{t+1}^{(n-1)} + \omega_{t+1}^{(n-1)}, \frac{M_{t+1}}{E_t[M_{t+1}]} \right] - E_t \left[ \omega_{t+1}^{(n-1)} \right] \\
&= \underbrace{-Cov_t \left[ r x_{t+1}^{(n-1)}, \frac{M_{t+1}}{E_t[M_{t+1}]} \right]}_{\text{log-risk premium}} - \underbrace{E_t \left[ \omega_{t+1}^{(n-1)} \right] - Cov_t \left[ \omega_{t+1}^{(n-1)}, \frac{M_{t+1}}{E_t[M_{t+1}]} \right]}_{\text{approximation error}}
\end{aligned} \tag{28}$$

We can decompose the approximation error  $\omega_{t+1}^{(n-1)}$  into a part that is correlated with the state variables  $\nu_{t+1}$ , and a part  $\xi_{t+1}^{(n-1)}$  that is orthogonal:

$$\omega_{t+1}^{(n-1)} = \phi^{(n-1)'} (W_0 + W_1 X_t) + \phi^{(n-1)'} \nu_{t+1} + \xi_{t+1}^{(n-1)}, \tag{29}$$

where we have assumed that the conditional mean of  $\omega_{t+1}^{(n-1)}$  is  $\phi^{(n-1)'} (W_0 + W_1 X_t)$ . Replacing this and the pricing kernel into the no arbitrage equation gives:

$$\begin{aligned}
E_t \left[ r x_{t+1}^{(n-1)} \right] &= -Cov_t \left[ r x_{t+1}^{(n-1)}, \frac{M_{t+1}}{E_t[M_{t+1}]} \right] - E_t \left[ \omega_{t+1}^{(n-1)} \right] - Cov_t \left[ \omega_{t+1}^{(n-1)}, \frac{M_{t+1}}{E_t[M_{t+1}]} \right] \\
&= Cov_t \left[ r x_{t+1}^{(n-1)} + \omega_{t+1}^{(n-1)}, \nu_{t+1} \right] \Sigma_t^{-1} (\lambda_0 + \lambda_1 X_t) - E_t \left[ \omega_{t+1}^{(n-1)} \right] \\
&= (\beta^{(n-1)'} + \phi^{(n-1)'}) (\lambda_0 + \lambda_1 X_t) - \phi^{(n-1)'} (W_0 + W_1 X_t).
\end{aligned} \tag{30}$$

Matching coefficients with (4) gives:

$$\begin{aligned}
\beta^{(n-1)'} \gamma_0 &= (\beta^{(n-1)'} + \phi^{(n-1)'}) \lambda_0 - \phi^{(n-1)'} W_0 \\
\beta^{(n-1)'} \gamma_1 &= (\beta^{(n-1)'} + \phi^{(n-1)'}) \lambda_1 - \phi^{(n-1)'} W_1
\end{aligned}$$

We can stack these equations to obtain:

$$(\beta' + \phi') \lambda_0 = \beta' \gamma_0 + \phi' W_0 \quad (31)$$

$$(\beta' + \phi') \lambda_1 = \beta' \gamma_1 + \phi' W_1 \quad (32)$$

So setting prices of risk to zero implies:

$$\lambda_0 = 0 \implies \gamma_0 = -(\beta\beta')^{-1} \beta\phi' W_0 \quad (33)$$

$$\lambda_1 = 0 \implies \gamma_1 = -(\beta\beta')^{-1} \beta\phi' W_1 \quad (34)$$

Because  $\omega$  is observable, the parameters  $\phi^{(n-1)'}$ ,  $W_0$ , and  $W_1$  can be estimated from (29) using the three step regression approach. In the Gaussian case,  $\phi' W_0 = \frac{1}{2}\beta^* S_0$  and  $\phi' W_1 = \frac{1}{2}\beta^* S_1$ .

## B Tables and Figures

Table 1: **Three Factor Model: Yield Pricing Errors**

This table summarizes the time series properties of the yield pricing errors  $\hat{u}$  implied by the three factor specification of our model. The sample period is 1986:01-2008:12. "mean", "std", "skew", and "kurt" refer to the sample mean, standard deviation, skewness, and kurtosis of the yield errors;  $\rho(1)$ ,  $\rho(6)$ , and  $\rho(12)$  denote the autocorrelation coefficients of order one, six, and twelve, respectively.

	n = 12	n = 24	n = 36	n = 60	n = 84	n = 120
mean	0.025	0.014	-0.002	-0.010	-0.001	-0.010
std	0.061	0.025	0.015	0.028	0.015	0.036
skew	0.062	1.008	-0.250	-0.833	-0.140	0.045
kurt	5.633	4.422	7.329	3.800	2.716	2.745
$\rho(1)$	0.881	0.858	0.786	0.867	0.888	0.856
$\rho(6)$	0.665	0.368	0.483	0.390	0.424	0.355
$\rho(12)$	0.377	0.034	0.140	0.076	0.164	0.104

Table 2: **Three Factor Model: Return Pricing Errors**

This table summarizes the time series properties of the return pricing errors  $\hat{e}$  implied by the three factor specification of our model. The sample period is 1986:01-2008:12. "mean", "std", "skew", and "kurt" refer to the sample mean, standard deviation, skewness, and kurtosis of the yield errors;  $\rho(1)$ ,  $\rho(6)$ , and  $\rho(12)$  denote the autocorrelation coefficients of order one, six, and twelve, respectively.

	n = 12	n = 24	n = 36	n = 60	n = 84	n = 120
mean	0.035	-0.006	-0.018	0.002	0.012	-0.017
std	0.433	0.353	0.291	0.734	0.531	2.129
skew	0.214	0.341	1.180	0.194	0.256	0.275
kurt	9.502	4.461	11.793	5.694	4.861	4.769
$\rho(1)$	0.143	0.111	0.014	0.080	0.174	0.092
$\rho(6)$	0.162	0.043	0.117	0.049	0.043	-0.003
$\rho(12)$	0.107	0.050	0.112	0.087	0.072	0.050
$\bar{R}^2$	0.972	0.997	0.999	0.998	0.999	0.995

**Table 3: Three Factor Model: Market Prices of Risk**

This table summarizes the estimates of the market price of risk parameters  $\gamma_0$  and  $\gamma_1$  for the three factor specification of our Approximate Affine Model.  $t$ -statistics are reported in brackets. The standard errors have been computed according to the bootstrap procedure laid out in Section 2.4.

	$\gamma_0$	$\gamma_{1.1}$	$\gamma_{1.2}$	$\gamma_{1.3}$
X1	0.035	-0.033	0.029	-0.021
t-stat	( 1.970)	(-1.725)	( 1.670)	(-1.380)
X2	0.037	-0.028	0.015	-0.013
t-stat	( 2.099)	(-1.116)	( 0.738)	(-0.769)
X3	0.004	-0.041	-0.002	-0.002
t-stat	( 0.083)	(-0.799)	(-0.046)	(-0.056)

**Table 4: Five Factor Model: Yield Pricing Errors**

This table summarizes the time series properties of the yield pricing errors  $\hat{u}$  implied by the five factor specification of our approximate affine model. The sample period is 1986:01-2008:12. "mean", "std", "skew", and "kurt" refer to the sample mean, standard deviation, skewness, and kurtosis of the yield errors;  $\rho(1)$ ,  $\rho(6)$ , and  $\rho(12)$  denote the autocorrelation coefficients of order one, six, and twelve, respectively.

	n = 12	n = 24	n = 36	n = 60	n = 84	n = 120
mean	0.003	0.007	0.005	0.005	0.007	0.004
std	0.024	0.009	0.006	0.006	0.005	0.008
skew	0.265	0.140	0.447	-0.344	0.672	-1.394
kurt	2.768	2.845	6.211	3.534	3.560	7.775
$\rho(1)$	0.916	0.912	0.789	0.915	0.843	0.827
$\rho(6)$	0.594	0.597	0.447	0.591	0.436	0.374
$\rho(12)$	0.271	0.441	0.153	0.364	0.272	0.156

Table 5: **Five Factor Model: Return Pricing Errors**

This table summarizes the time series properties of the return pricing errors  $\hat{e}$  implied by the five factor specification of our approximate affine model. The sample period is 1986:01-2008:12. "mean", "std", "skew", and "kurt" refer to the sample mean, standard deviation, skewness, and kurtosis of the yield errors;  $\rho(1)$ ,  $\rho(6)$ , and  $\rho(12)$  denote the autocorrelation coefficients of order one, six, and twelve, respectively.

	n = 12	n = 24	n = 36	n = 60	n = 84	n = 120
mean	-0.000	0.000	-0.001	0.001	-0.000	-0.000
std	0.042	0.054	0.109	0.056	0.190	0.514
skew	0.059	0.110	0.466	0.411	-0.119	0.233
kurt	18.583	5.381	12.987	6.950	5.748	5.906
$\rho(1)$	-0.201	0.093	-0.090	0.159	-0.030	0.045
$\rho(6)$	0.240	0.075	0.185	-0.017	0.047	0.044
$\rho(12)$	0.014	0.170	0.055	0.040	0.112	0.117
$R^2$	1.000	1.000	1.000	1.000	1.000	1.000

Table 6: **Five Factor Model: Pseudo Market Prices of Risk**

This table summarizes the estimates of the quasi market price of risk parameters  $\gamma_0$  and  $\gamma_1$  for the five factor specification of our model.  $t$ -statistics are reported in brackets. The standard errors have been computed according to the bootstrap procedure laid out in Section 2.4.

	$\gamma_0$	$\gamma_{1.1}$	$\gamma_{1.2}$	$\gamma_{1.3}$	$\gamma_{1.4}$	$\gamma_{1.5}$
X1	0.035	0.001	0.009	-0.002	0.018	-0.029
t-stat	( 2.127)	( 0.030)	( 0.477)	(-0.122)	( 1.516)	(-2.145)
X2	0.040	0.005	-0.008	0.010	0.021	-0.030
t-stat	( 2.611)	( 0.215)	(-0.440)	( 0.627)	( 1.606)	(-2.256)
X3	0.022	-0.158	0.040	-0.041	-0.166	0.025
t-stat	( 0.495)	(-2.300)	( 0.759)	(-0.984)	(-3.837)	( 0.548)
X4	-0.054	0.080	0.032	-0.035	0.153	0.083
t-stat	(-0.610)	( 0.591)	( 0.288)	(-0.452)	( 1.421)	( 1.042)
X5	0.004	0.240	-0.069	0.070	0.140	-0.160
t-stat	( 0.050)	( 1.845)	(-0.740)	( 0.861)	( 1.169)	(-2.033)



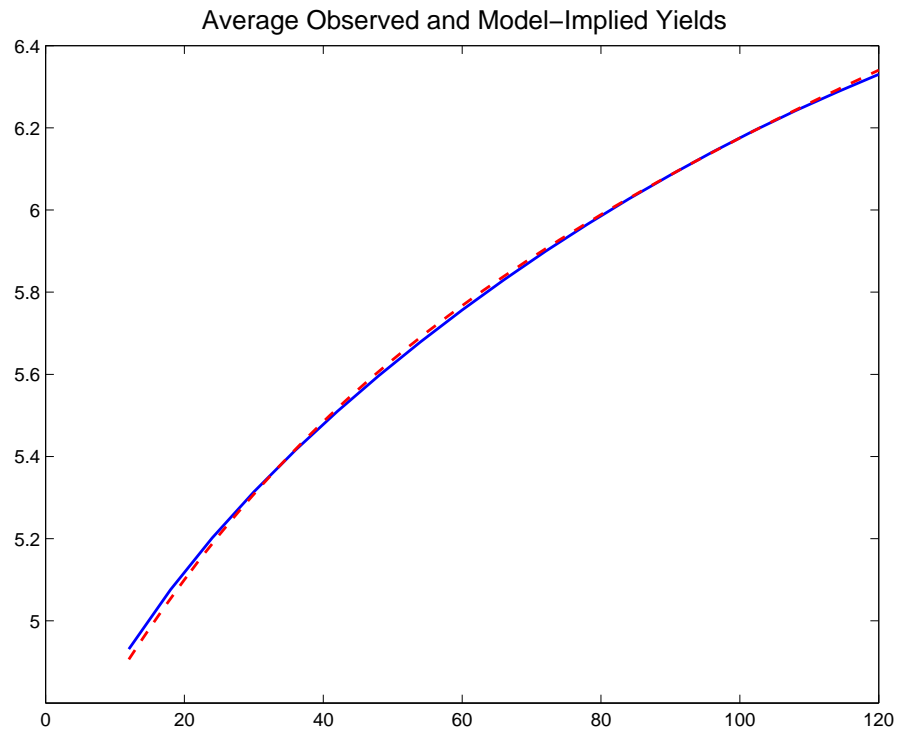
Table 7: **RMSEs from Approximate Affine Specifications and Nelson-Siegel Model Relative to RW: 2003:01-2008:12**

This table summarizes the root mean squared forecast errors relative to those implied by a random walk for yields. The models have been estimated using data from 1986:01 until the period when the forecast is made. The forecasting period is 2003:01-2008:12. “AA(3)” and “AA(5)” refer to the three factor and five factor specifications of our Approximate Affine Model, “AA(5)+CP” denotes the five factor specification augmented by the Cochrane-Piazzesi return forecasting factor, and “NS” denotes the Diebold-Li (2006) version of the three-factor Nelson-Siegel model with VAR factor dynamics.

	n = 12	n = 24	n = 36	n = 60	n = 84	n = 120
<b>Panel : 1 months ahead</b>						
<i>AA(3)</i>	0.935	0.987	0.994	1.006	1.003	1.040
<i>AA(5)</i>	0.890	0.942	0.974	0.991	0.994	1.014
<i>AA(5) + CP</i>	0.889	0.925	0.955	0.984	0.999	1.020
NS	1.116	1.030	1.030	1.071	1.047	1.033
<b>Panel : 6 months ahead</b>						
<i>AA(3)</i>	0.940	0.991	0.999	1.000	1.017	1.065
<i>AA(5)</i>	0.868	0.914	0.945	0.987	1.022	1.074
<i>AA(5) + CP</i>	0.848	0.892	0.925	0.973	1.014	1.072
NS	1.002	0.994	0.996	1.025	1.032	0.995
<b>Panel : 12 months ahead</b>						
<i>AA(3)</i>	0.936	0.968	0.976	0.991	1.039	1.145
<i>AA(5)</i>	0.777	0.795	0.816	0.870	0.952	1.103
<i>AA(5) + CP</i>	0.746	0.759	0.782	0.845	0.942	1.114
NS	0.922	0.927	0.940	0.999	1.064	1.090

Figure 1: **Three Factor Model: Average Observed and Model-Implied Yields**

This figure plots average observed yields against those implied by the Three Factor Approximate Affine Model.



**Figure 2: Three Factor Model: Standard Deviation of Observed and Model-Implied Yields**

This figure plots the unconditional standard deviations of observed yields against those implied by the Three Factor Approximate Affine Model.

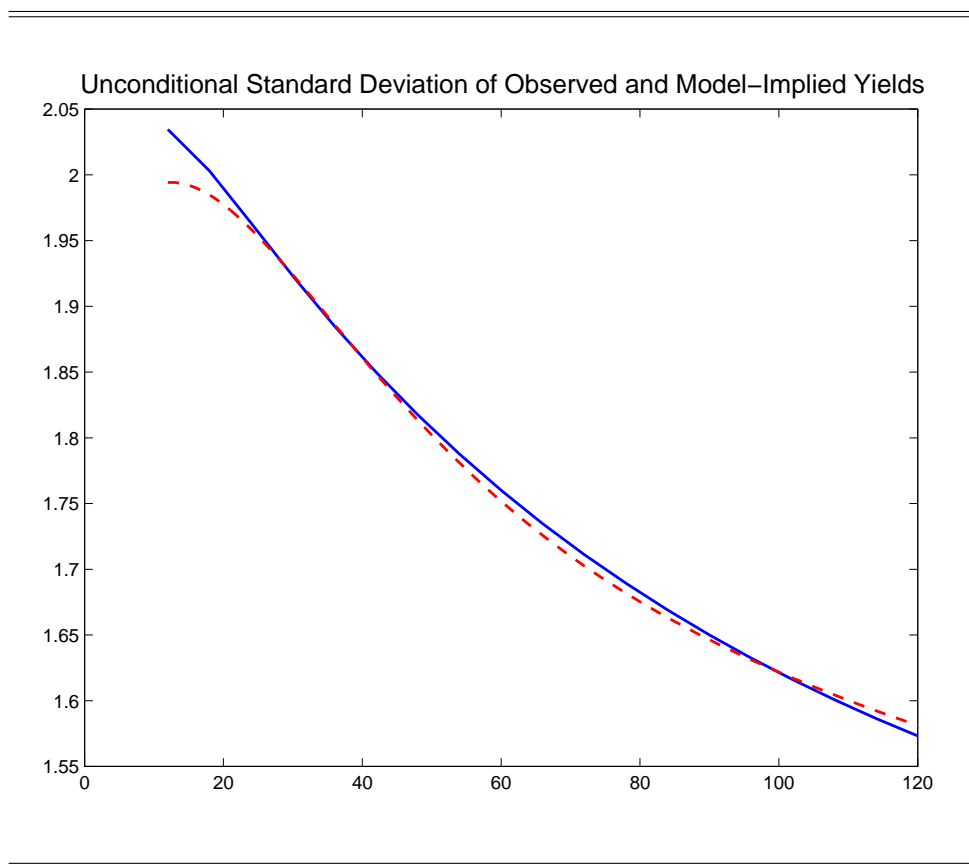
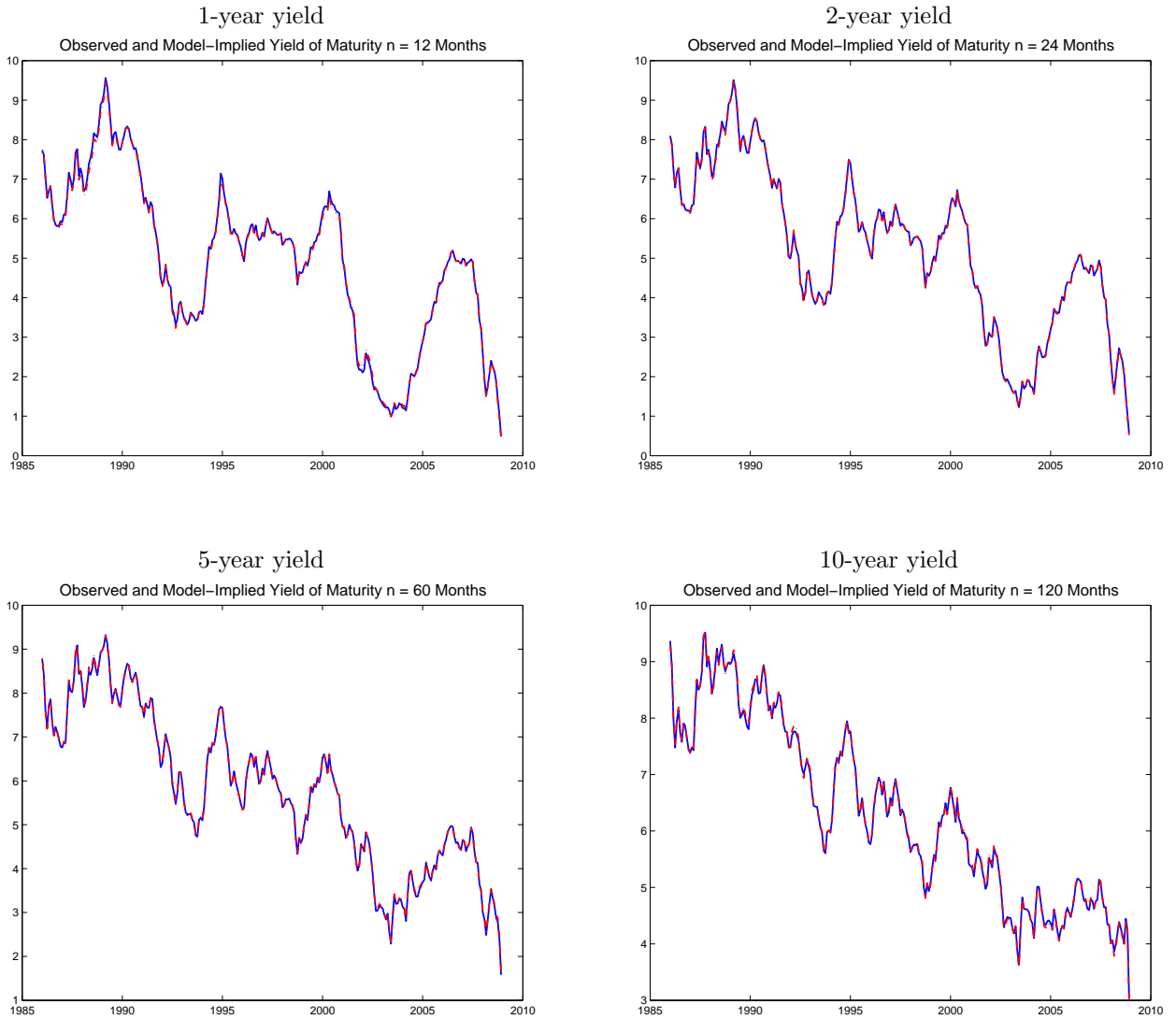


Figure 3: **Three Factor Model: Observed and Model-Implied Yields**

This figure provides plots of the observed and fitted yields for the 1-year, 2-year, 5- and 10-year maturities. The observed yields are plotted by solid lines, whereas dashed lines correspond to yields implied by the three factor model.



**Figure 4: Three Factor Model: Observed and Model-Implied Excess One-Month Holding Returns**

This figure provides plots of the observed and model-implied excess one-month holding returns for the 1-year, 2-year, 5- and 10-year maturities. The observed returns are plotted by solid lines, whereas dashed green lines correspond to actual model-implied returns, and dash-dotted red lines to model-implied expected one-month excess holding returns.

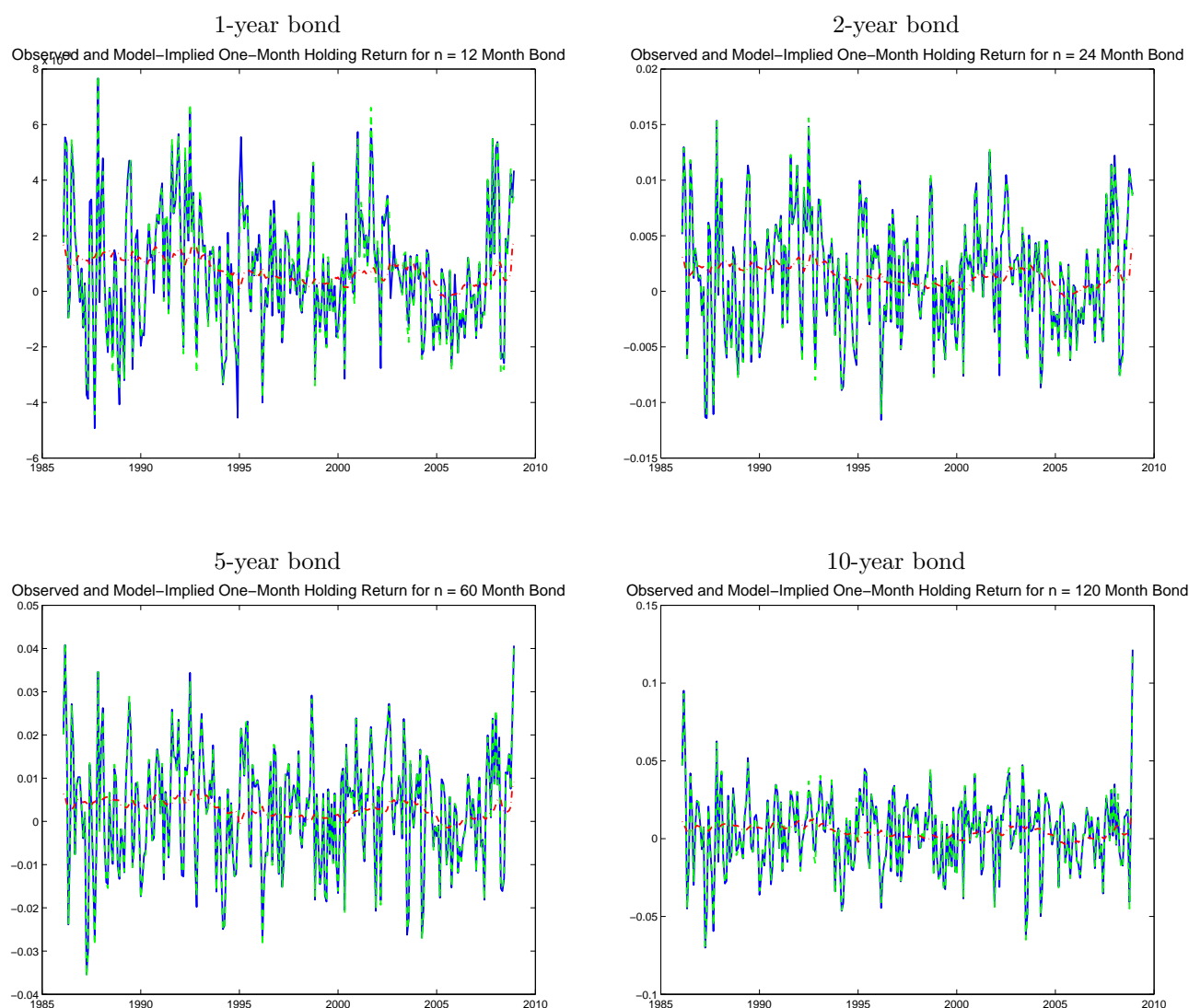


Figure 5: **Three Factor Model: Factor Loadings for Yields**

This figure provides a plot of the yield loadings  $-\frac{1}{n}B_n$  implied by the Three Factor Approximate Affine Model. The coefficients can be interpreted as the response of the  $n$ -month yield to a contemporary shock to the respective factor.

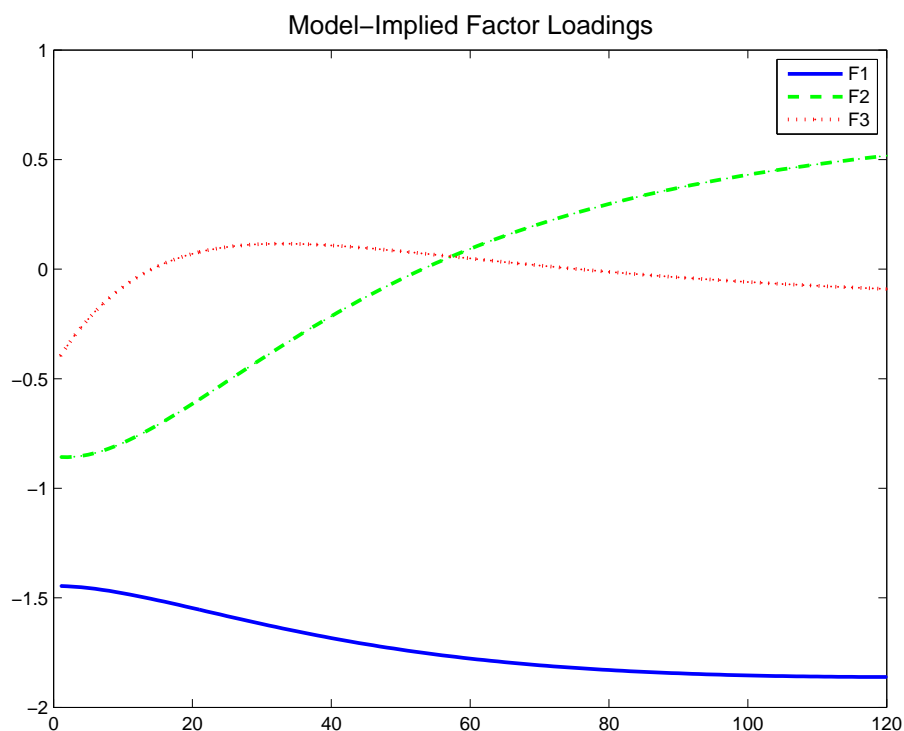


Figure 6: **Three Factor Model: Factor Loadings for Expected Returns**

This figure provides a plot of the expected return loadings  $B'_n \gamma_1$  implied by the Three Factor Approximate Affine Model. The coefficients can be interpreted as the response of the expected one-month excess holding return on an  $n$ -month bond to a contemporary shock to the respective factor.

---

---

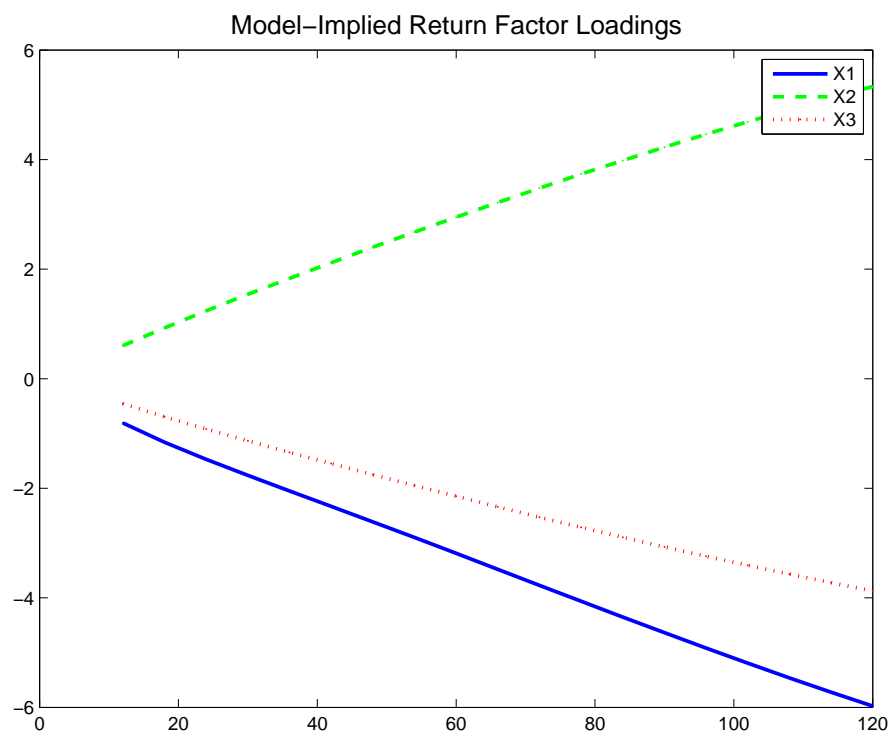
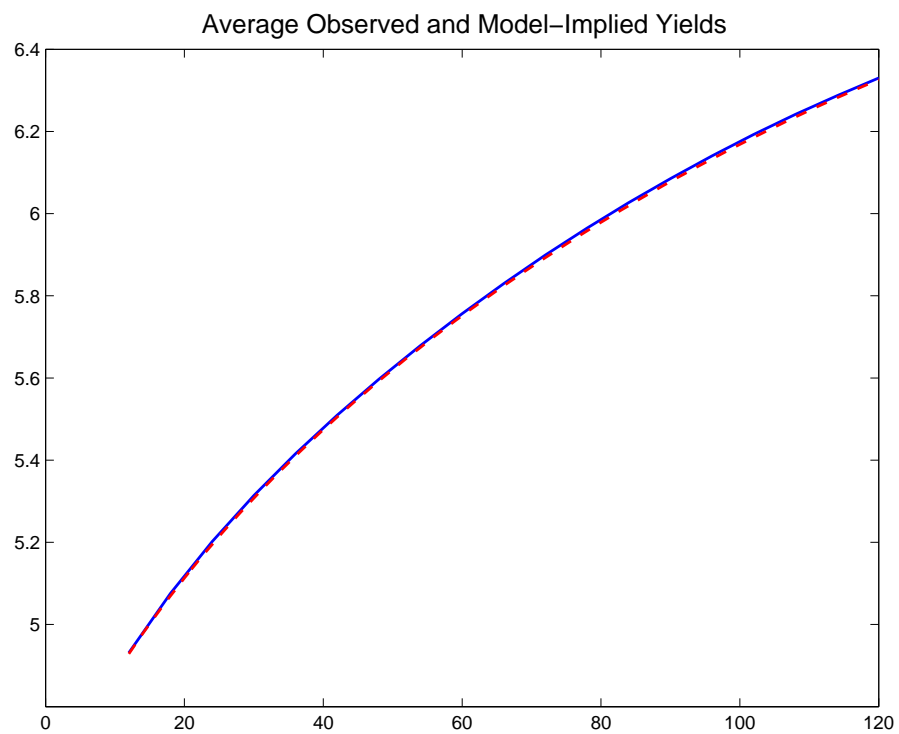


Figure 7: **Five Factor Model: Average Observed and Model-Implied Yields**

This figure plots average observed yields against those implied by the Five Factor Approximate Affine Model.





**Figure 8: Five Factor Model: Standard Deviation of Observed and Model-Implied Yields**

This figure plots the unconditional standard deviations of observed yields against those implied by the Five Factor Approximate Affine Model.

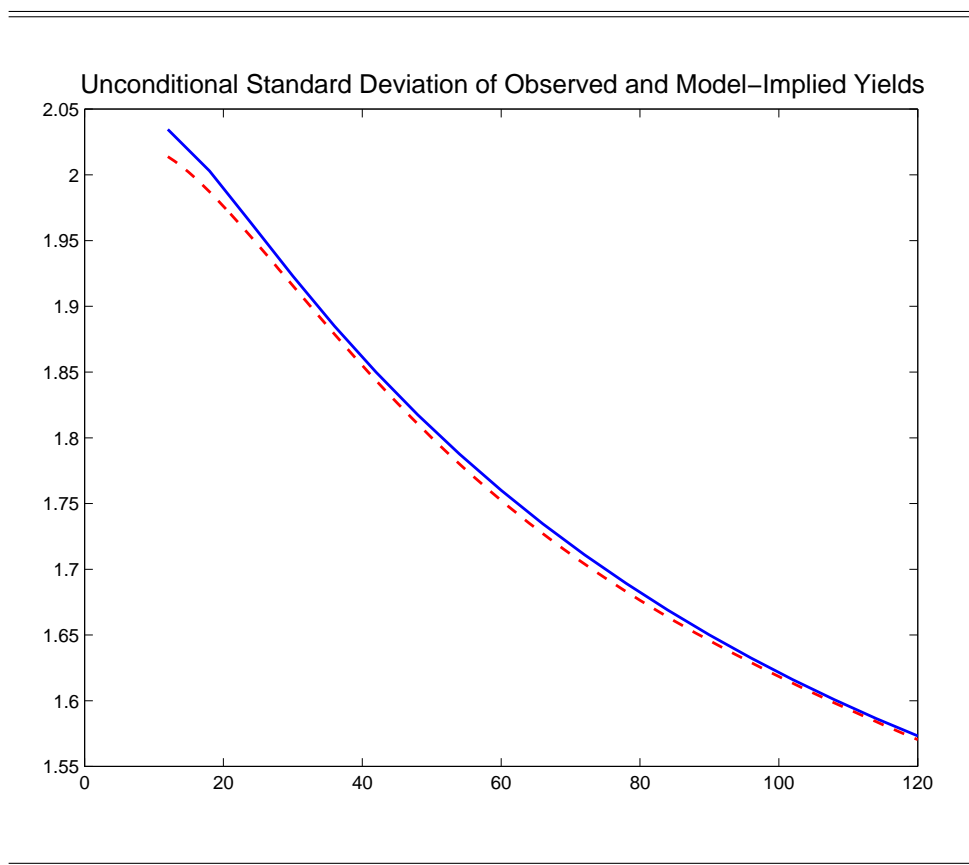
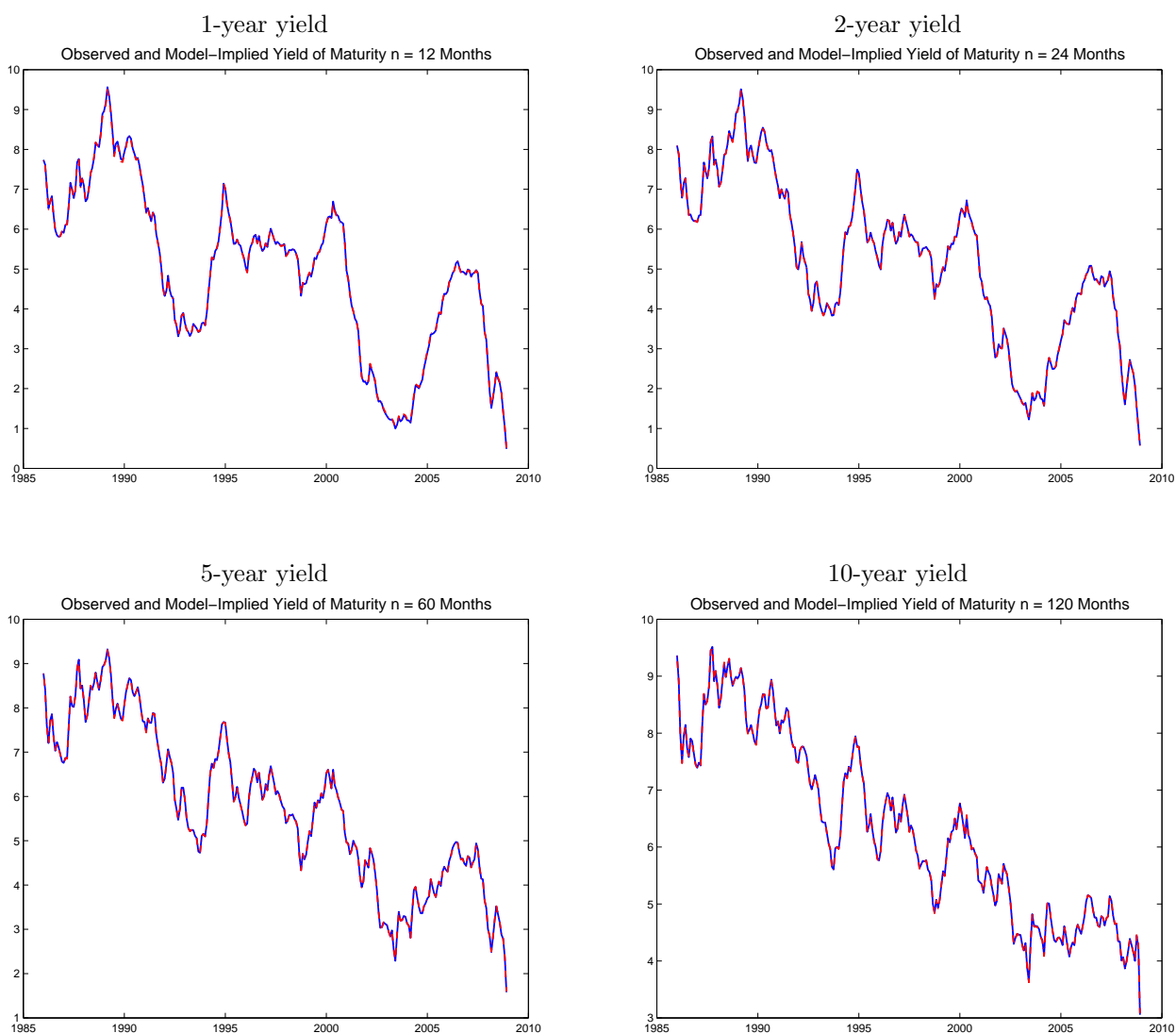


Figure 9: **Five Factor Model: Observed and Model-Implied Yields**

This figure provides plots of the observed and fitted yields for the 1-year, 2-year, 5- and 10-year maturities. The observed yields are plotted by solid lines, whereas dashed lines correspond to yields implied by the three factor model.



**Figure 10: Five Factor Model: Observed and Model-Implied Excess One-Month Holding Returns**

This figure provides plots of the observed and model-implied excess one-month holding returns for the 1-year, 2-year, 5- and 10-year maturities. The observed returns are plotted by solid lines, whereas dashed green lines correspond to actual model-implied returns, and dash-dotted red lines to model-implied expected one-month excess holding returns.

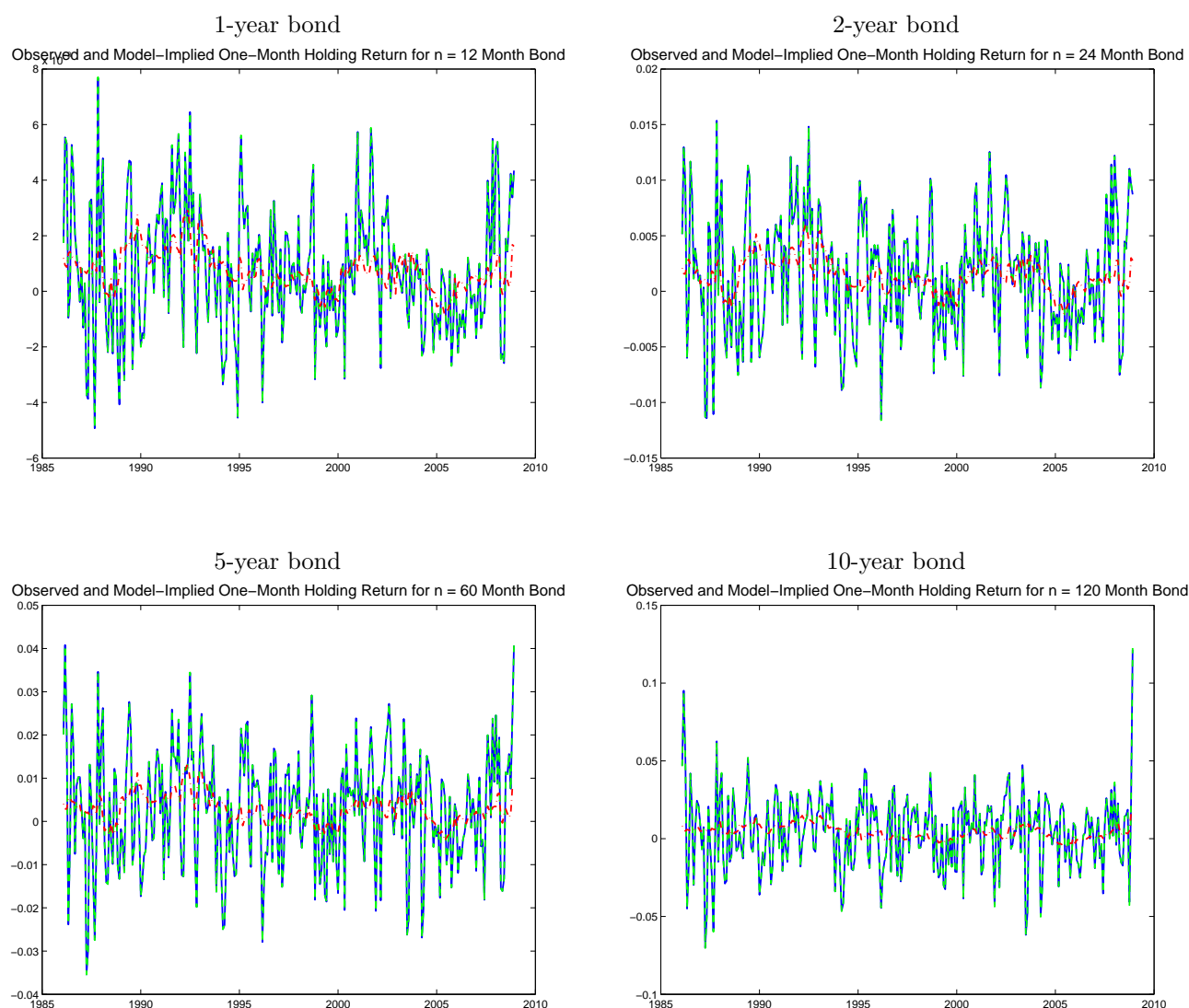


Figure 11: **Five Factor Model: Factor Loadings for Yields**

This figure provides a plot of the yield loadings  $-\frac{1}{n}B_n$  implied by the Five Factor Approximate Affine Model. The coefficients can be interpreted as the response of the  $n$ -month yield to a contemporary shock to the respective factor.

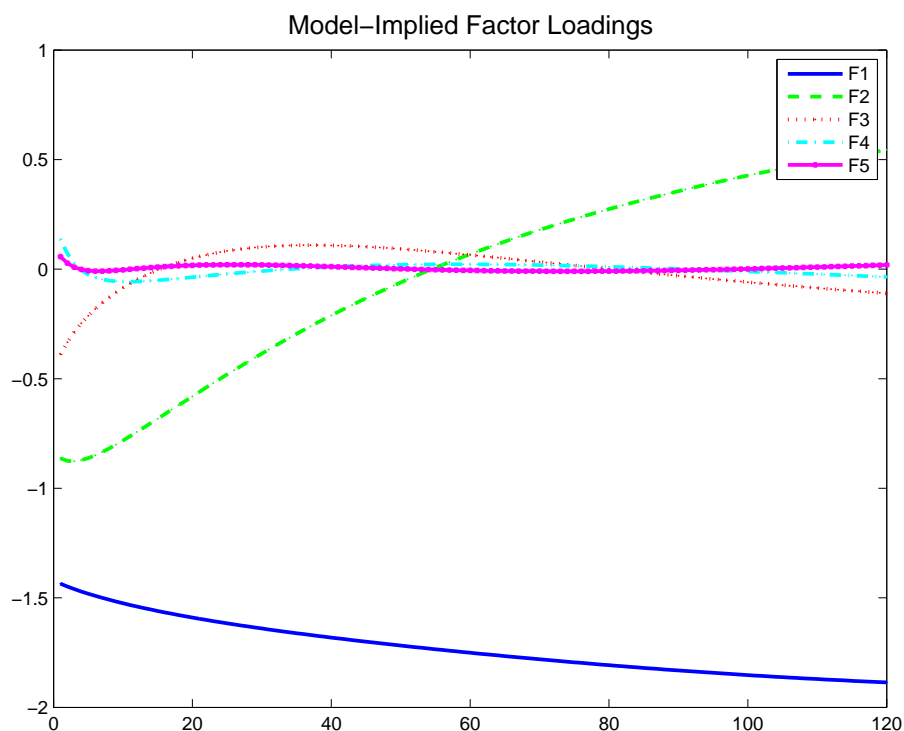


Figure 12: **Five Factor Model: Factor Loadings for Expected Returns**

This figure provides a plot of the expected return loadings  $B'_n \gamma_1$  implied by the Five Factor Approximate Affine Model. The coefficients can be interpreted as the response of the expected one-month excess holding return on an  $n$ -month bond to a contemporary shock to the respective factor.

---

---

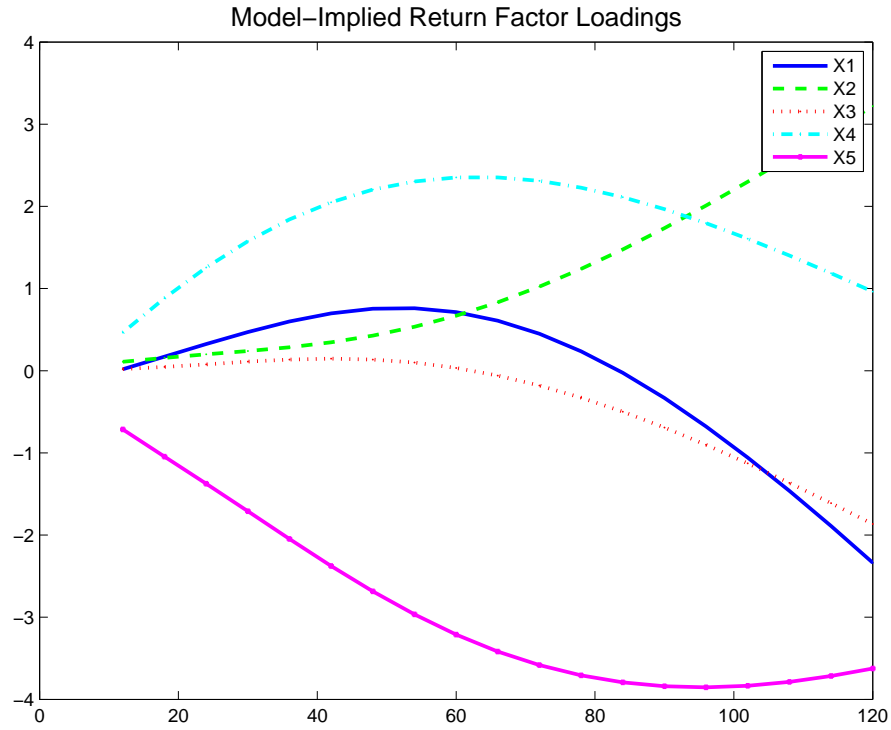


Figure 13: **Five Factor Model: Regression Coefficients  $\beta^{(n)}$  versus recursive pricing parameters  $B_n$**

This figure provides plots of the coefficients  $\beta^{(n)}$  from the regression (5) of log excess holding period returns on the state variable innovations versus the recursive pricing parameters  $B_n$  given in (12b). The red diamonds represent the former while the blue solid line corresponds to the latter.

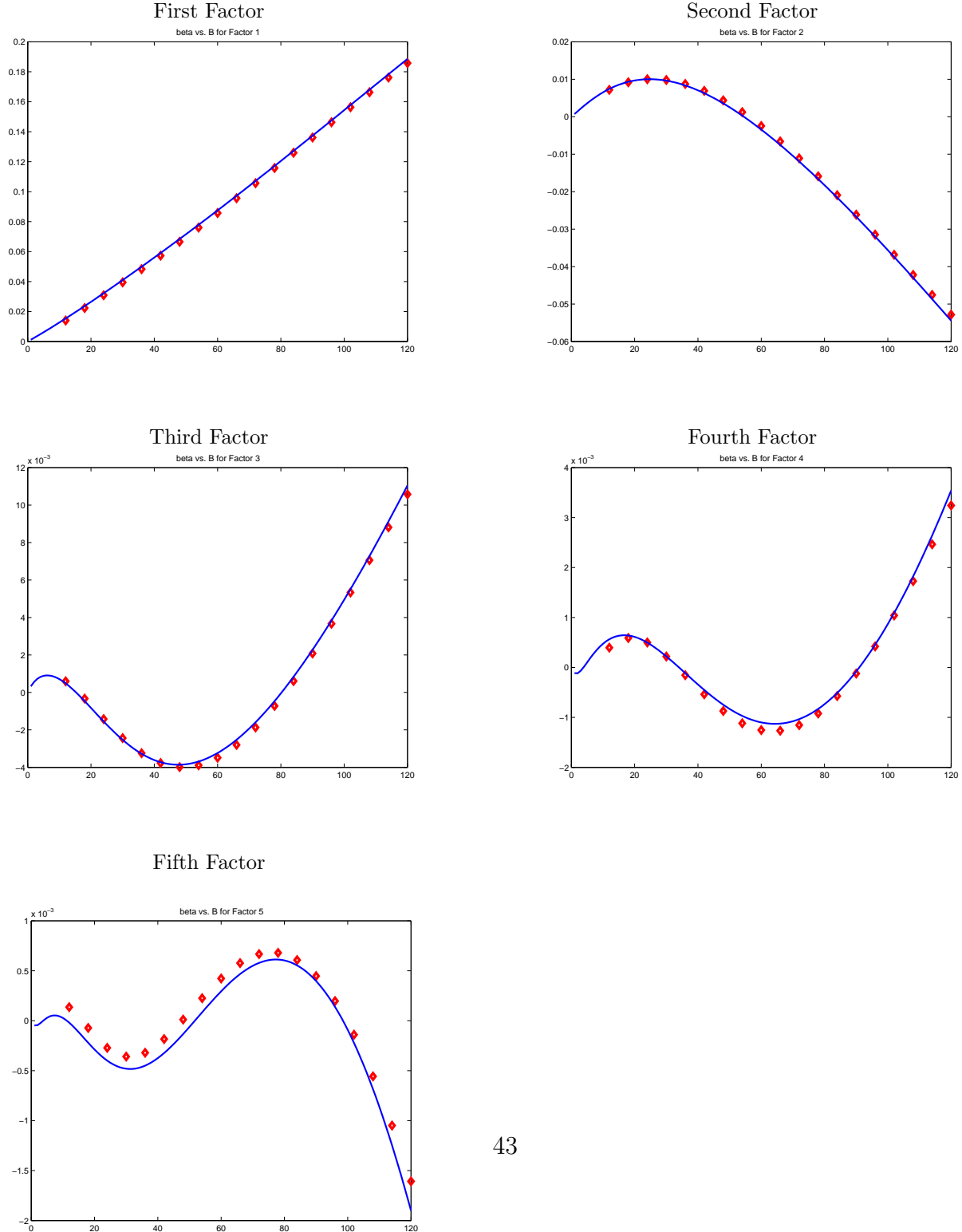
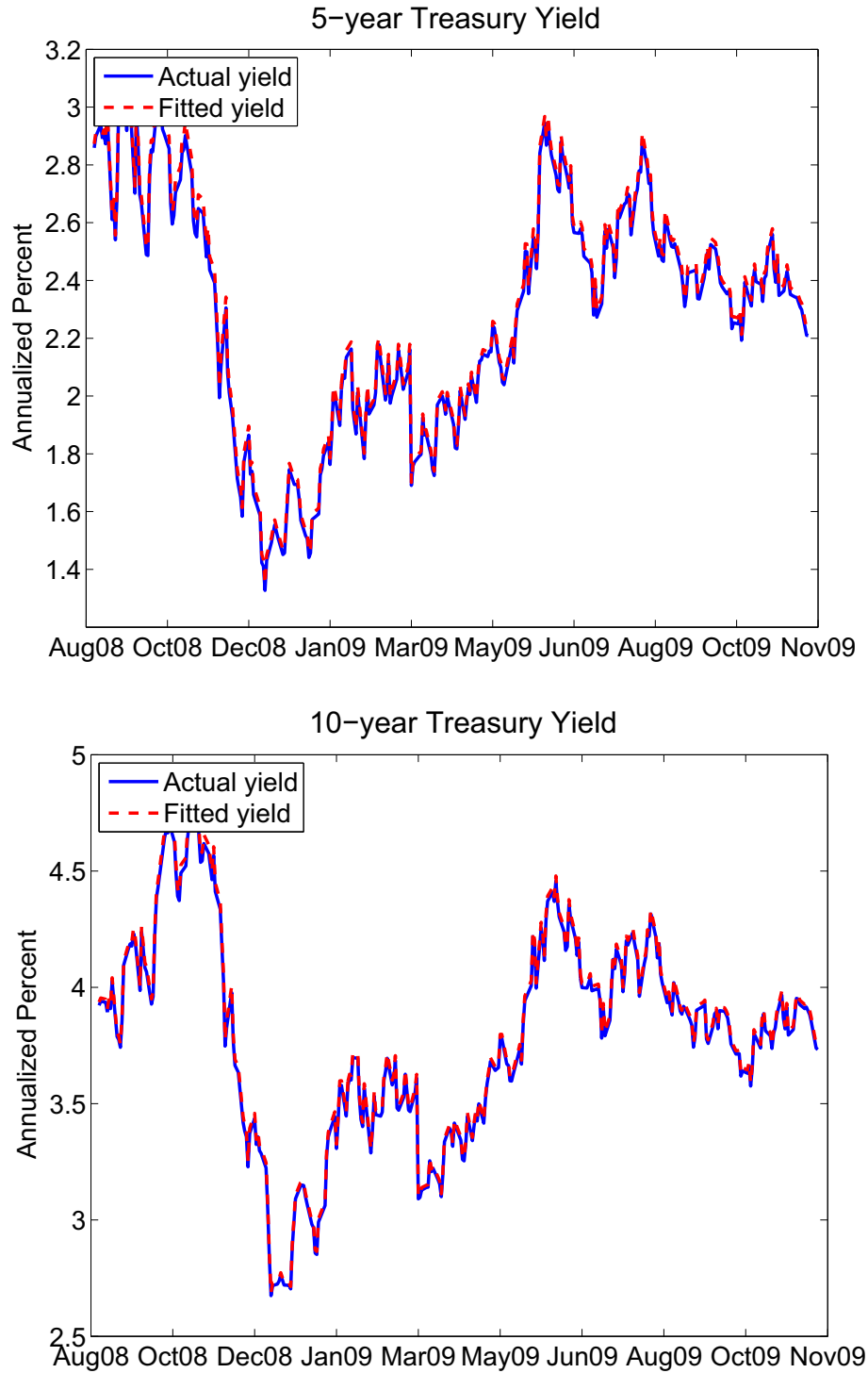


Figure 14: Five Factor Model: Observed and Model-Implied Daily Yields



---

Figure 15: 10-Year Treasury Term Premium and 1-Month MOVE Index

---

